ilmedia

TECHNISCHE UNIVERSITÄT ILMENAU

Chiang, Feng-Kuang; Knauf, Rainer; Tso, Tai-Cheng :

Analyzing the students' achievements of Taiwan web-based mathematics competition by data mining

Zuerst erschienen in:
E-Learn 2008 : World Conference on E-Learning in Corporate, Government, Healthcare & Higher Education ; November 17 - 21, 2008, Las Vegas, Nevada, USA / AACE, Association for the Advancement of Computing in Education. Ed. by Curtis J. Bonk ... ; Vol. 4. - Chesapeake, VA : Assoc. for the Advancement of Computing in Education (AACE), 2008. - S. 2560-2565

Analyzing the Students' Achievements of Taiwan Web-based Mathematics Competition by Data Mining

Feng-Kuang Chiang
Dept. of Industrial Technology Education
National Kaohsiung Normal University
Taiwan (ROC)
forite@gmail.com

Rainer Knauf
Artificial Intelligence Group, Faculty of Computer Science and Automation
Ilmenau University of Technology
Germany
Rainer.Knauf@tu-ilmenau.de

Tai-Cheng Tso
Dept. of Mathematics,
National Kaohsiung Normal University
Taiwan (ROC)
taicheng@nknucc.nknu.edu.tw

Abstract: Experts from three different fields will cooperate to explore this study, including e-learning experts from Taiwan; data mining experts from Canada, and Artificial Intelligence experts from Germany. Through this new analysis technique (data mining), new discoveries can be concluded in the field of web-based mathematics competition. Researchers try to explore these relations between different students' portfolios and achievements of Taiwan Web-based Mathematics Competition. These findings will provide valuable information for instructors, researchers and government. In particular, optimal learning conditions in- and outside the learning scenario can be explored.

Background

More and more technologies like mobile-learning, e-library or virtual learning environment are cropping up around the world and challenging traditional teaching materials; anyone can get information and share knowledge through the WWW easily. The rapid development of the everywhere-accessible Internet has caused certain changes to instructional forms (Lee, Yeh, Kung, & Hsu, 2007). Harold (2006) pointed out the integration of new information technologies into the society in general and into the educational environment in particular seems destined to increase. In the age of Internet and information, the education is moving toward to a teaching and learning evolution. Today's Information Technology (IT) provides easy access to mathematic learning content for all students. In addition, the way of competition coheres with the principle of interest and the self-learning in teaching method arousing the interests and attentions of the competitors.

Taiwan participated in many international competitions with an outstanding performance such as International Mathematical Olympiad (IMO), Asian Pacific Mathematics Olympiad (APMO), and Po Leung Kuk Primary Mathematics World Competition (PMWC). However, not everyone has the opportunity to join the competitions. It is the authority of specific students or phenom. Furthermore, the entries for IMO and APMO are high school students, even though there some competitions are held in Japan, Indonesia, India, Mainland China (including Hong Kong) for primary school, and junior high. These are still unavailable to every student. If we hold an activity for any grades, it would greaten the opportunities for every student to communicate on mathematic. As a consequence, this year, 2008.9~2009.3, a general web-based competition was opened for K1-K12 student through internet in Taiwan.

Purpose of the study

The study mainly design and develop a Web-based Mathematics Competition System and analyze the relation between the students' competitive achievements and their learning conditions and attitudes by data mining. Researchers will collect the students' portfolio, including three parts (1. Social information, 2. Educational information, 3.Others) and explore the difference of students' competitive achievements in 2008 web-based mathematics competition. As a result, (1) a success estimation in advance to the competition can be provided by applying the data mining result to new data and, more important, (2) optimal learning conditions in- and outside the learning scenario can be explored.

Significance of the study

Data mining in relation to enterprise resource planning is the statistical and logical analysis of large sets of transaction data, looking for patterns that can aid decision making (Monk & Wagner, 2006). It is the same in educational areas; researchers can obtain an in-depth understanding of the current educational situation and social phenomenon, through the means of analyzing a large number of data. The application of Data Mining technologies in learning processes has been proven to be a successful means for planning educational processes such as a university study (Knauf et al., 2008, Sakurai et al., 2008).

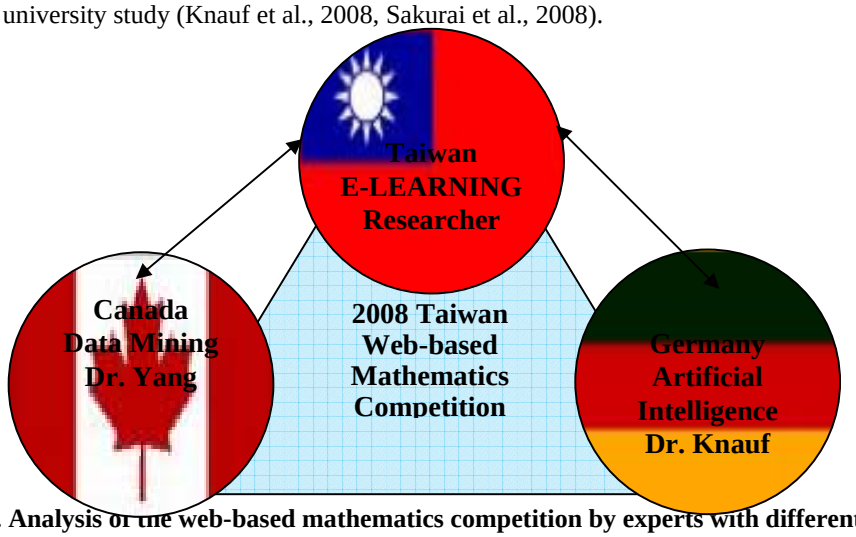

Figure 1. Analysis of the web-based mathematics competition by experts with different scientific backgrounds

In the proposed study, experts from three different fields and countries will cooperate to explore the above mentioned objectives, including e-learning experts from Taiwan; data mining experts from Canada, and Artificial Intelligence experts from Germany (See Figure 1).

Through this new analysis technique (data mining), new discoveries can be concluded in the field of web-based mathematics competition in particular, but also of learning conditions for mathematical content in general. The expected findings will provide valuable information for instructors, researchers and government. These insights may be used for optimizing learning conditions of pupils as well as for estimating particular talents or weaknesses, when mathematical skills will be challenged. Based on such insights, potentially talented pupils may be encouraged to rise to mathematical challenges such as a competition or even a university study in this field.

As a long term objective of this a subsequent studies such results may be useful for an optimal human resource management and recruiting processes.

Methodology

Framework of the study

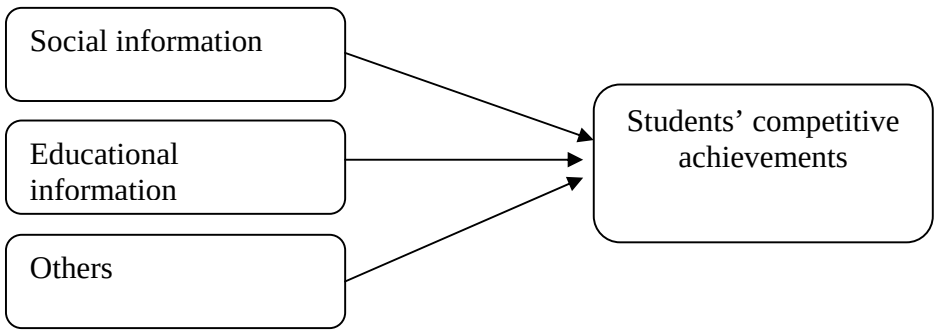

Figure2. The framework of research design

Figure 2 illustrates the framework of the research design. There are three different learning stages (elementary, junior, and senior high school) students who took part in this Web-based Mathematics Competition in Taiwan. Researchers attempted to explore the relation between three portfolios and students' competitive achievements.

Data collection

This study was to conduct an online survey, and participators can fill out register information before they login the competitive system. Participators come from different learning stages students, including public-private elementary school, junior high school, and senior high school. The date will be collected form 2008.10-2009.2 for four months.

Data mining and analysis

Data mining is the process of sorting through large amounts of data and picking out relevant information. Figure 3 shows typical steps of a data mining process. It is usually used by business intelligence organizations, and financial analysts, but is increasingly being used in the sciences to extract information from the enormous data sets generated by modern experimental and observational methods. It has been described as "the nontrivial extraction of implicit, previously unknown, and potentially useful information from data" (Frawley, Piatetsky-Shapiro, & Metheus, 1992) and "the science of extracting useful information from large data sets or databases" (Hand, Mannila & Smyth, 2001).

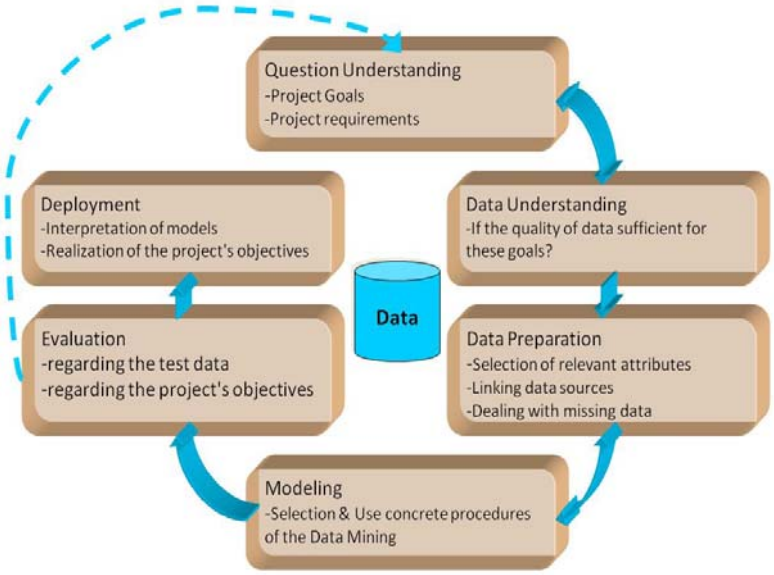

Figure 3. The process of the data mining model

In particular, we analyze the correlation between various items that may influence the students' learning capability in general and their talents for mathematics in particular. For this purpose, we plan the following working steps:

1. **Development of a catalog of properties, which may have an influence on the students' learning capabilities**
   This catalog should contain all issues of the pupils' life including
   a. Social information such as family background, grow-up area (city/village) and others,
   b. Educational information such as school type, favorite subjects and others, and
   c. Other potential influence factors such as sport activities, food patterns and others.
   A very first outline of such a catalog is shown in appendix A.
2. **Development of an evaluation scale of values for the above mentioned attributes**
3. **Analysis of common data mining technologies for their applicability to the present problem**

Here, we consider the most common technologies Top Down Induction of Decision Trees (TDIDT) such as ID3 and C 4.5 and their extensions (towards numeric values, uncertainty and missing values) for their applicability to the problem. If no one is really appropriate, we plan to develop an own, application specific technology similar to the one in (Knauf et al., 2008).

4. **Application of the selected or developed technology to the mathematic competition**
   Here, we develop an
   a. appropriate methodology to acquire all necessary data,
   b. a program to perform the data mining on it, and
   c. an appropriate presentation of the results

   In particular, by using the concept of Information Entropy, the "most informative" property (as found in step one above) will be identified as the route of the decision tree under construction. Most informative, in this context, means providing maximum information on the issue, how successful, i.e. up to which level (see next section) the competition will be passed. According to the number of this property's values (or value ranges), that appear in the data, the entire set of students' data will be partitioned into the same number of subsets, which will be handled accordingly for constructing each sub-tree. This process will continue with the sub-sub-trees and so on as long as the subsets still contain students' data with different success levels.

Web-based math competition system

Figure 4 shows web-based math competition system. The process of students took part in web-based MCS as follows; first, students have to register personal information on this website, and then they can logon the web-based MCS to exam the four different levels mathematical questions. The Web-based MCS will provide the feedback and record the process of how students answer these questions. Once participators answer the wrong for four times, they will be logged off by the system automatically. Figure 5 shows the Web-based Mathematic Competitive website. The figure 6 is the Web-based Mathematic Competitive system webpage.

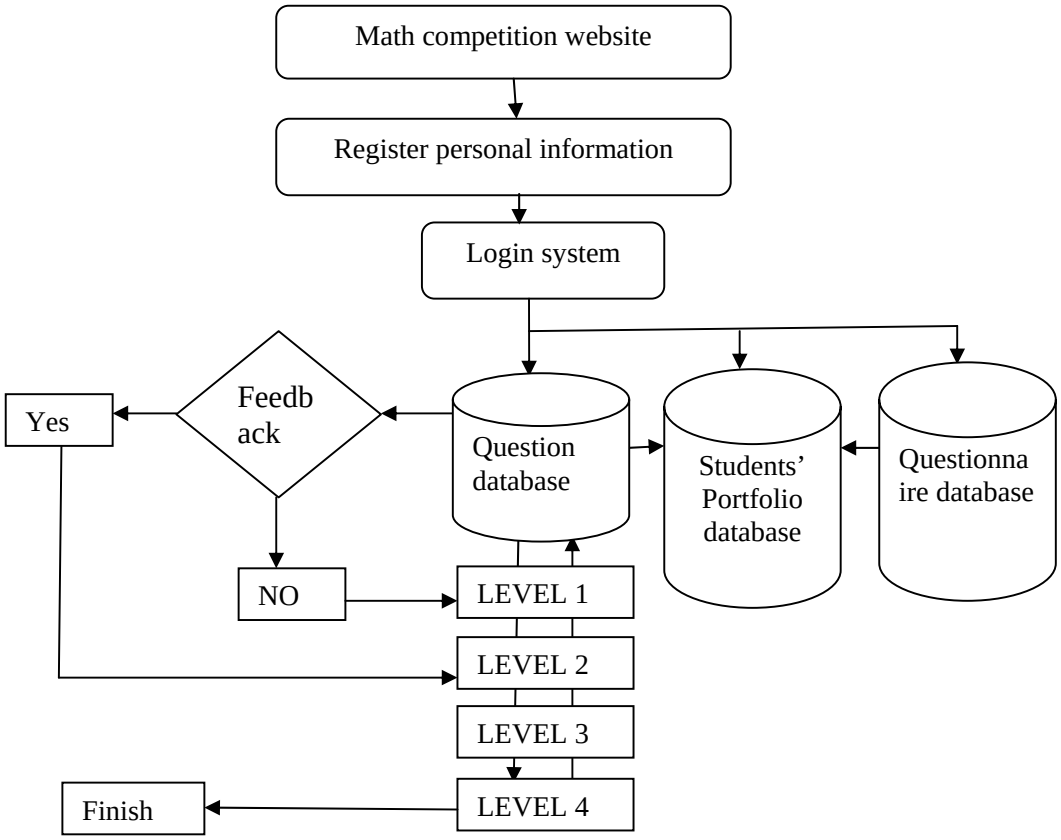

Figure 4. Web-based math competition system

Figure 5. Web-based Mathematic Competitive website

Figure 6. Web-based Mathematic Competitive system webpage

Anticipated Tasks and Results

1. To explore the relative literature review
2. To design and development the 2008 Web-based Mathematical Competition System
3. To impalement the 2008 Web-based Mathematical Competition activities in Taiwan
4. To analyze these relations between different Students' Portfolios and achievements of Taiwan Web-based Mathematics Competition by data mining in the way sketched in section 2.3
5. To finish this study and publish papers in international journal

References

Frawley, W., Poatetsky-Shapiro, G. & Metheus, C. (1992). Knowledge Discovery in Databases: An Overview," AI Magazine,13(3), 57-50.

Harold, R. (2006). *Information Technology integration in higher education: A case study of historically black university.* Unpublished doctoral dissertation, Capella University.

Hand, D., Mannila, H. & Smyth, P. (2001).Principles of Data Mining. MIT Press, Cambridge, MA. ISBN 0-262-08290-X.

Knauf, R., Böck, R., Sakurai, Y., Tsuruta, S. (2008): A Priori Evaluation and Refinement of Curricula by Data Mining over Storyboards.Proc. of the 21th International Florida Artificial Intelligence Research Society Conference 2008 (FLAIRS-2008), Miami, FL, USA, Menlo Park, CA: AAAI Press, ISBN 978-1-57735-365-2, pp. 335-340, 2008.

Lee, C. H., Yeh, D., Kung R. J., & Hsu, C. S. (2007). The Influences of Learning Portfolios and Attitudes on Learning Effects in Blended E-learning for Mathematic. *Journal of Educational Computing Research. 37(4).* 331-250.

Monk, E. & Wagner, B.(2006). Concepts in Enterprise Resource Planning, Second Edition. Thomson Course Technology, Boston, MA. ISBN 0-619-21663-8.

Sakurai, Y., Dohi, S., Tsuruta, S., Knauf, R. (2008): Modeling Academic Education Processes by Dynamic Storyboarding. Accepted Paper. Journal of Educational Technology & Society, vol. 11, ISSN 1436-4522 (online) and 1176-3647 (print). International Forum of Educational Technology & Society (IFETS), 2008

Appendix A: Students' Portfolio

A. Social information

1. Name:
2. Gender: boy girl
3. ID card number:
4. Birth date: year month day
5. TEL:
6. ADD:
7. E-MAIL:
8. How many brother and sister are you? ___elder brother ___elder sister ___younger brother ___younger sister
9. How old is your father: (1)No father (2)21-30 (3)31-40 (4)41-50 (5) 51-60 (6)61-70 (7)over 71 years old
10. How old is your mother? (1)No mother (2)21-30 (3)31-40 (4)41-50 (5) 51-60 (6)61-70 (7)over 71year old
11. What kind of occupation is your father? NO job,_____job
12. What kind of occupation is your mother? NO job, _____job

B. Educational information

1. Which group do you participate in? (1) Elementary school (2) Junior high school (3) Senior high school
2. Who is your mathematical instructor?
3. Which school do you study in?
4. Are you a Gifted Education class student? (1)Yes (2)No
5. How long do you use computers? (1)one year (2)2 years (3) 3 years (4) 4 years (5) over 5 years
6. How long do you use the internet? (1)one year (2)2 years (3) 3 years (4) 4 years (5) over 5 years
7. How many classes are in your school? (1)1-10 classes (2)11-20 classes (3) 21-30 classes (4) 31-40classes (5) 41-50 classes (6) over 51 classes
8. Do you study in cram school now? (1)Yes (2)No
9. Do you take mathematical course in cram school now? (1)Yes (2)No
10. What is your Favorite subject?
11. What subject do you dislike most?
12. How many hours do you study math after school every day? (1)0 (2)1 (3)2 (4)3 (5)4 (6)over 5 hours
13. Do you feel overstrained or underchallenged by the learning requirements of your school? (1) underchallenged (2) partly underchallenged (3) neither nor (4) partly oversrained (5) overstrained

C. Others

1. Are you a handicap on your body? (1)Yes (2)No
2. Do you belong to any religion? (1)Yes (2)No
3. Do you make sports or do exercise at least once a week? (1)Yes (2)No
4. Do you like to play computer games? (1)Yes (2)No
5. Generally, are you pleased with you current circumstances of life? (1) yes (2) rather yes (3) partly yes/partly no (4) rather no (5) no
6. According the future of your life, are you optimistic? (1) yes (2) rather yes (3) partly yes/partly no (4) rather no (5) no